

Durham Research Online

Deposited in DRO:

14 January 2020

Version of attached file:

Published Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Beckmann, J.F. (2019) 'Heigh-ho : CPS and the seven questions – some thoughts on contemporary complex problem solving research.', Journal of dynamic decision making, 5 . p. 12.

Further information on publisher's website:

<https://doi.org/10.11588/jddm.2019.1.69301>

Publisher's copyright statement:

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Heigh-Ho: CPS and the seven questions – some thoughts on contemporary Complex Problem Solving research

Jens F. Beckmann¹

¹Durham University, United Kingdom.

In this paper, I share some reflections on a set of questions posed by the editors of Journal for Dynamic Decision Making in relation to research in complex problem solving (CPS). These questions, especially in their combination, suggest problems in CPS research with regard to identity, direction, and purpose. I focus on three issues. The first issue is the diversity in objectives and methodological approaches subsumed under the label of CPS. The resulting conceptual ambiguity makes it challenging to define CPS and thus, to identify ways to develop CPS research further. The second issue is the tendency in contemporary CPS research to employ psychometrics for autotelic purposes rather than utilising it as the tool that helps linking the conceptual with the empirical in psychological research. The third issue refers to the conceptual vacuum around the essential element of CPS, namely the concept of complexity. The tendency to substitute complexity (as a psychological concept) with difficulty (as a psychometric concept) tends to perpetuate the existing conceptual limitations and to compound the circularity associated with an operational definition of CPS. Indifference towards these issues is major hindrance to resolving the issues around identity, direction and purpose, and to making meaningful progress in CPS research.

Heigh-Ho: CPS and the seven questions – some thoughts on contemporary Complex Problem Solving research

As Henry Louis Mencken could have said, for every complex problem there is an easy solution that is neat, plausible, and wrong¹. I think it will not come to anyone's surprise that this rings also true for research in complex problem solving (CPS). Therefore, I shall not attempt to offer neat answers to the seven questions about CPS as they have been posed by the editors of this special issue. In the well-known fairy tale, each of the seven dwarfs asked a question when they – to their surprise – encountered the presence of Snow White in their abode after returning from a hard day's work in the mine. Whilst their questions led to one unifying answer, I do not expect this to be the case for the CPS-related questions asked here. Also, these CPS-related questions do not come to anyone's surprise, I am afraid:

(Q1) Why should there be CPS research? (Q2) How does current CPS research contribute to solving “real life” problems? (Q3) Can laboratory-based CPS research be rele-

vant? (Q4) What is the role of knowledge in CPS? (Q5) What is the role of strategies in CPS? (Q6) What is the role of intuition in CPS? (Q7) Do experts solve CPS problems differently to laypersons?

Asking questions is one way of stock-taking, allowing to reflect on what has been achieved, what is the status quo, and where should we go from here. Somewhat paradoxically, part of the answers to those questions lies in the questions themselves. From these questions it seems apparent that CPS research does have problems. These problems, some of them more complex than others, suggest issues with identity, direction and purpose. It is my impression that many of these issues are largely self-inflicted, which nurtures my optimism that resolving these is in our hands, or stated more appropriately, in our minds. The distinction between hands and minds shall also serve as reminder of the importance of the substantive, or the conceptual, and of the role the empirical has as a means to an end in the process of conducting empirical research in the social sciences. It therefore provides some orientation for how to address the problems complex problem solving has to solve.

Diversity

To get anywhere near a meaningful starting point for reflections on the above-stated questions, we first need to determine what CPS stands for. As it turns out, this is a bigger problem as anticipated by some. The label “CPS” has been used in many, quite different contexts and for diverse purposes. For instance, CPS has been used for labelling a research paradigm employed to study psychological concepts such as information processing, learning, decision-making, causal reasoning, knowledge acquisition (see Q4), strategy use (see Q5), and more. CPS has also been used as a label for a predominantly cognitive ability (i.e., “complex problem solving ability”). Another, innocuous but also rather uninformative use of CPS would be as a label for a class of observed behaviour that individuals exhibit when confronted with computer-simulated microworlds. These are just a few examples for the diversity in the use of the label CPS (i.e., as description of a research methodology, as description of latent psychological constructs, or as description of observable behaviour). Such diversity and the resulting lack of clarity in meaning creates a considerable

¹ The exact quote is “... there is always a well-known solution to every human problem – neat, plausible, and wrong” (Mencken, 1921, p. 158).

challenge to passing a verdict on CPS research's *raison d'être* (see Q1). This diversity problem will not necessarily be solved by focussing on only one use and ignoring the other alternative uses. By looking at the body of empirical research on CPS related to abilities one might get the impression that depending on the correlations found to scores from other measures of cognitive abilities (e.g., reasoning tests), CPS is "narrowed or downgraded" to a skill, or "widened or upgraded" to a competency, or anything in between. Such conceptual flexibility is a side effect of a predominantly operational definition of CPS (i.e., CPS is what one measures with tasks that carry the label CPS). Treating ability, skill and competency as synonyms, as it happens occasionally, makes the problem of vagueness even worse. Or, as they say, what happens in vagueness stays in vagueness. There are, of course, attempts to get a more conceptual handle on the definition problem. In a more recent attempt to navigate the complex landscape of CPS research, Dörner and Funke (2017) have suggested that the term complex problem solving should be reserved for dealings with ill-defined problems. There is some irony in it as CPS research appears rather ill-defined itself. This seems to create an odd impasse: Any efforts to define CPS would make it potentially non-CPS. The situation, however, is a bit more complex than that. The distinction between well-defined and ill-defined is not a dichotomy, which creates the challenge of determining where well-defined ends and ill-defined begins on an imaginary definition scale. In addition, problems that were previously considered ill-defined might become ever so slightly better defined as we progress in our conceptual understanding. The fact that we tend to contemplate the same or similar questions over the past 4 decades seem to suggest that the pace with which we make conceptual progress is rather slow. Whilst this might reduce the risk of dealing with moving targets when it comes to defining CPS, it should not be mistaken as justification for contend with the status quo. An other way of addressing the ambiguity problem and to cater towards a more or less peaceful co-existence of different conceptual foci and methodological approaches in CPS research would be to declare CPS a mere "umbrella" term. Emphasising communalities (e.g., shared interests, goals, and tool use) and downplaying (real or otherwise) differences has been necessary and strategically functional in the early days of CPS research when the primary challenge was to establish a "new idea". Having reached a state now where everyone seems to have staked their claims in this mine (heigh-ho, heigh-ho!) the all-encompassing umbrella has outlived its usefulness. One side effect of continuing to subsume considerable diversity under a common label is that constructive impulses (e.g., via feedback from outside or within CPS research) tend to have limited productive impact as it seems to be always "the others" who should take notice. It is not too difficult to see how that can stifle conceptual progress in a research area such as CPS.

A non-definition of an ill-defined problem

In seeming contradiction to the above-mentioned notion of "true" CPS being ill-defined, I continue with an attempt to offer some clarifications regarding the term complex problem solving. I start at the end: problem solving. A simple, yet powerful and therefore widely accepted framework of problem solving posits that all problems are comprised of three components: a current state of affairs; a goal state of affairs; and a set of steps that need to be taken to move

from one to the other. I would argue that this applies even to ill-defined problems. Problems vary in ways of which these components are specified or known, and which of these components are expected to be identified by the problem solver. That means, the actual problem could be either to identify the set of actions needed to reach a goal state, or to find out what the unknown end state is going to be, or to "backward engineer" what the initial state of affairs was given the end state and the set of transition rules. For example, in a performance management context, having received feedback that current performance levels are deemed unsatisfactory (i.e., evaluative framing of the current state of affairs) and being made aware of what the expected performance levels are (i.e., defining the goal state of affairs) leaves one with the problem of identifying how to move from one to the other. Or, in the aftermath of a financial collapse of a company one might be tasked with identifying "where it all went wrong". The problem solving focus here would be on the transition processes from retrospectively inferred previous states of affairs that have led to the current state of affairs. Another example might be to estimate the likelihood of convalescence (i.e., goal state of affairs) given the clinical symptoms presented in a patient (i.e., current state of affairs) and the knowledge about the effect principles of a specific combination of treatments (i.e., set of probabilistic transitions). Based on such rather simple framework using three basic problem components and varying what is known and what is unknown has the potential to capture a wide range of (even "real-life", see Q2) problems. Problems, of course, also differ with regard to the level of detail with which these components are or can be specified. "Real-life" problems tend to be more towards the lower end of this spectrum and might therefore be considered ill-defined. Here the effort required in clarifying the current state of affairs and / or the desired state of affairs can be much greater than figuring out the actual steps that need to be taken to move from one to the other (Wood, Cugin, & Beckmann, 2009). When striving for real life relevancy it is worth to keep in mind that not all "real-life" problems are necessarily "ill-defined" and being "ill-defined" does not make a problem a "real-life" problem. Real life relevance (see Q3) or "ecological validity" (as one of those terms afflicted by inflationary use) of CPS is neither "proven" via correlations with some "real-life" variables (see Q2), nor achieved by "making it look like the real thing". The use of semantically enriched cover stories and accordingly labelled variables in CPS tasks can be seen as one example for such attempts. As it has been shown (e.g., Beckmann & Goode, 2014, 2017) this has more often than not detrimental effects on the quality of what is measured in the end. As Borsboom and colleagues remind us "... the problem of validity cannot be solved by psychometric techniques or models alone. On the contrary, it must be addressed by substantive theory. Validity is the one problem in testing that psychology cannot contract out to methodology" (Borsboom, Mellenbergh, & Van Heerden, 2004, p. 1062). Reflections in relation to Q2 could benefit from a better conceptual foundation. One impulse to that effect could come from Lewin's notion of *Geschehenstyp* (or "principle of the type of event", as an imperfect translation, Lewin, 1927/1992). It posits that the task paradigm used in experimental, laboratory-based research needs to reflect the structure and function of the processes involved in the class of real-life situations that is targeted (see Q3). It is the "common logic of research in the laboratory and the field" what generalisation claims have to be

based on (Gadenne, 1976). In the context of CPS research this means that cognitive (or otherwise) processes triggered by experimental tasks need to overlap with what we expect to take place when dealing with said real-life problems. As said, this cannot be achieved through attempts to create some form of superficial resemblance of surface features (such as variable labels) between lab task and targeted real-life problem. It means, however, that some *ex ante* ideas are needed about both the real-life problem and the laboratory task (see Q3). Such ideas have to be derived from theory based task analyses. Consulting what psychological theories have to offer (truly) prior to engaging in data collection reduces the risk of declaring *post hoc* interpretations of correlation patterns as “proof of a theory”. Another advantage of putting the conceptual horse back in front of the empirical carriage is that it will help us to identify where gaps exist in our theories. Meaningful research should focus on strengthening the theoretical foundations of our conceptual understanding of the psychological processes involved when people attempt to deal with complex, dynamically changing challenges in their lives. The empirical side of research is a means to that end. Consulting “the oracle of numbers” instead cannot serve as a substitute for the conceptual work we ought to be doing.

The elephant in the ... problem space

After some clarifications regarding problem solving, I will now focus on the remaining component of the term CPS, which is “complex”. As has been discussed previously at lengths and in detail (Beckmann & Goode, 2017; see also Beckmann, Birney & Goode, 2017; Beckmann, 2010; Birney, Beckmann, & Seah, 2016), I argue that the lack of a shared understanding of what we mean by complexity in CPS research is the major barrier to any meaningful progress. This problem is rooted in a predominantly data-driven take on CPS, which promotes the tendency to put the (empirical) cart before the (conceptual) horse. In research that starts with data and ends with data the concept of complexity is substituted by the notion of difficulty. Difficulty, however, is a psychometric concept with limited conceptual and explanatory reach. It descriptively informs us about that one challenge (e.g., an intelligence test item or a CPS task) has been successfully tackled by fewer people than another challenge. The former is then declared being more difficult than the latter. If one were interested in the reasons for such outcome (i.e., an explanation), one would be left with a tautological argument (i.e., because more people solved the second problem). Simply replacing difficulty with complexity, as it happens notoriously often in CPS research, will not provide an explanation. As stated previously “... complexity reflects *ex ante* considerations of the cognitive demands imposed by the task and the circumstances under which the task is to be performed ... , which makes complexity a primarily cognitive concept. Difficulty, in contrast, is experiential, person-bound and by definition, statistical” (Beckmann, Birney, & Goode, 2017, p. 2). By using difficulty and complexity interchangeably one mistakes descriptions for explanations.

We need to understand, or description \neq explanation

Some research where certain result patterns are interpreted as indications of intuition (see Q6) or strategy (see Q5)

could serve as examples for such tendency. The term “intuition” seems all too readily be used as an “explanation” for situations where perceived success in problem solving is seemingly independent of structural, causal, strategic, or prior “world-knowledge” (see Q4). Such constellation of knowledge-free success could, however, be an effect of employing ineffective measures of problem-relevant knowledge, which often is a result of a weak conceptual foundation of the operationalisation of knowledge in CPS. There are, of course, other potential reasons for not finding correlational links between knowledge and problem solving performance. For instance, when problem solvers are asked to bring a system’s current state to a set goal state, they might follow the simple, knowledge free heuristic of an intervention-by-intervention optimisation (e.g., Beckmann & Goode, 2017; Beckmann & Guthke, 1995). Labelling such interaction behaviour “intuition” would be misleading. An uncritical use of “intuition” in CPS creates the considerable risk of masking remaining limitations in our conceptualisation (i.e., underpinning theories) and operationalisation (i.e., the variables derived from measures), and the links between the two. The underlying problem is that many labels we tend to use for describing observed phenomena in CPS carry pseudo-explanatory meanings, so that descriptions are being mistaken as explanations. This also applies to the use of “strategy” (see Q5) for labelling clusters of interaction patterns that have been identified by statistical routines trawling sets of secondary data (a.k.a. log file analysis). Again, “intuition” would be an attractive candidate for labelling interaction behaviour that has resulted in acceptable outcomes, yet does not exceed the statistically necessary systematicity threshold for constituting a “strategy” cluster. In other words, failing to identify some sort of systematicity pattern (or residual, in statistical terms) in the average problem solver’s interaction with the task, which would otherwise be given the label “strategy X”, is not necessarily a reason for labelling it “intuition”. In the context of strategy-focussed CPS research that starts with *a priori* considerations of how effective and efficient interaction behaviour for knowledge acquisition should look like, the so-called VOTAT² strategy plays a prominent role. VOTAT, however, has little relevance in dealing with real world problems as complex, dynamic (“ill-defined”?) problems faced “in the wild” as such problems tend to not afford one with the freedom to vary none or only one variable at a time. Hence, strategies used by successful problem solvers in the lab will not necessarily separate the successful from the less successful problem solvers in real life (see Q2, Q3, & Q5). Using expert-novice comparisons as another example for a research topic within CPS (see Q7), conceptual groundwork needs to clarify a few basics before engaging in data collection (or the analysis of existing data). For instance: Given that expertise is domain and knowledge specific, and given that “true” complex problem solving is “ill-defined” and/or “ill-structured”, what CPS expertise would look like? The answer to this question has implications for how to measure CPS expertise. Another issue that should be addressed conceptually prior to engaging in the empirical

² VOTAT, which stands for vary-one-variable-at-a-time, is only functional if such interaction behaviour is preceded by a zero intervention, which is needed to identify autonomic changes in the to be explored system. Therefore, the “desired” strategy should be more appropriately labelled VONAT (Vary-One-or-None-At-a-Time, Beckmann, Birney, & Goode, 2017, p. 4; Beckmann and Goode, 2014; p. 279; Beckmann and Goode, 2017, p. 9).

would be: Does it take 10 years of deliberate practice to become an expert in CPS? If so, what would be considered deliberate practice in CPS? Or, do the artificial time horizons as they are typical for computerised CPS tasks or simulations allow for a more time efficient acquisition of expertise? Does (extensive) experience in dealing with CPS tasks as they are used in CPS research satisfy the criterion of expertise? Given the diversity of approaches to measure CPS, predictions (i.e., made prior to “knowing the outcome”) of where expertise is expected to shine and where not would be needed. Otherwise one is confronted with a situation where those who outperform others are *post hoc* declared experts, which, of course, will have limited exploratory value. Unexpected (or undesired?) correlation patterns can be dealt with by either expressing doubts regarding the status of expertise achieved by problem solvers involved in the study, or by referring to the limited accessibility of experts’ knowledge – be it declarative or strategic (see Q4) – due to their higher levels of cognitive automatization, which, consequently, might even lead to the “conclusion” that experts come to problem solutions intuitively (see Q6). *Post hoc* interpretations such as these are readily available to “protect” a potentially inadequate conceptual foundation from being challenged (and eventually improved).

Structure in the ill-structured

Research in CPS (as is the case for any empirical research in the social sciences) can be characterised as a hierarchy of three objectives. These are description, explanation and intervention. “... a proper description of the phenomena of interest is a necessary (but not sufficient) precondition for developing adequate levels of understanding of the causal mechanisms that underlie them. An adequate understanding or explanation of the phenomena under question is, again, a necessary, yet not sufficient precondition for research to have meaningful impact in the “real world,” for example, in form of effective interventions.” (Beckmann, 2018, p. 121). Each objective calls upon specific sets of methodologies, including research design and sampling. For instance, the analysis of secondary observational data – as it is the foundation of log file analyses – would represent a mismatch to ambitions of establishing an understanding of causal effects. Or, if we were to consult the seven dwarfs in the well-known fairy tale: While going about their mining work enthusiastically the short folks sing “We dig dig dig ... up everything in sight; We dig up diamonds by the score; A thousand rubies, sometimes more; But we don’t know what we dig ’em for; We dig dig dig a-dig dig ...”. I am inclined to interpret these lyrics as a reminder of the potential risks that lie in data mining. Based on descriptive data we cannot make (*post hoc*) conceptual, explanatory claims. Descriptions of observed effects should not be interpreted as evidence of understanding. To avoid misunderstandings, log file analyses are certainly useful approaches to describe problem solvers’ (average) behaviour interacting with computerised CPS tasks. Results of log-file analyses can be effective in forming preliminary speculations about cognitive, conative or other psychological processes involved. Such speculations, translated in hypotheses, however, need to be properly tested in controlled experiments before we can claim to have gained some conceptual insights regarding their role in CPS.

Conclusion = Solution?

In this paper I have shared thoughts regarding some issues in CPS research that were innervated by a set of seven questions raised by the editors of this special issue. I wish these reflections to be perceived as an encouragement to a reversion to the conceptual foundations of problem solving and decision making. If CPS research were to do so, I suspect two things to emerge. On the one hand, I expect a realisation that we already know more than our current research practices seem to suggest. On the other hand, however, we will notice that our theoretical foundation has substantial limitations in describing, explaining and prescribing real world problem solving. Our efforts should primarily be directed towards the development or refinement of problem solving theories. Deficits in theory cannot be compensated for, by, say, larger sets of data. Such necessary reorientation would have to start with research geared towards a thorough description of the phenomena of interest (but without misinterpreting this as explanation), which then creates the foundation for research that aims to establish an understanding of these phenomena. The methodology necessary for that differs considerably from the one suitable for the purpose of description. Only if subsequent research efforts to prescribe interventions or pedagogies are based on such understanding, will we be in a promising position to devise effective strategies for the development and acquisition of problem solving competencies that enable individuals or groups of individuals to tackle the complex challenges in real life. This stricture also applies to using CPS tasks as assessment tools. Their meaningfulness and usefulness cannot be established by primarily relying on correlation matrices, it rather requires sufficient levels of conceptual understanding first. The linchpin of such reprioritisation of research efforts in CPS is a solid conceptual understanding of complexity. I started with the expectation that the seven questions stated at the beginning cannot be addressed by one single answer (as it was the case for the questions asked by the seven dwarfs as they tried to establish who has broken into their house). I would like to argue, however, that these seven CPS-related questions should be addressed by a single *question*: Why do we not have a better understanding of what complexity means in complex problem solving? If we continue to fail addressing this question, we are likely to mull over the same seven questions stated in the beginning of this paper in, say, ten years’ time, should CPS research not already have fallen into disregard by then. As it was a stumble that brought Snow White back to life in the fairy tale, I hope that an overdue step change in CPS research, namely to start with theory and not with data, will breathe fresh life into it too.

Declaration of conflicting interests: The author declares he has no conflict of interests.

Author contributions: The author is completely responsible for the content of this manuscript.

Handling editor: Andreas Fischer and Wolfgang Schoppek

Copyright: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Citation: Beckmann, J.F. (2019). Heigh-Ho: CPS and the seven questions – some thoughts on contemporary Complex Problem Solving research *Journal of Dynamic Decision Making*, 5, 12. doi: 10.11588/jddm.2019.1.69301

Published: 31 Dec 2019

References

- Beckmann, J.F. (2010). Taming a beast of burden – On some issues with the conceptualization and operationalisation of cognitive load. *Learning and Instruction*, 20, 250-264. doi:10.1016/j.learninstruc.2009.02.024
- Beckmann, J.F., & Goode, N. (2014). The benefit of being naïve and knowing it: The unfavourable impact of perceived context familiarity on learning in complex problem solving tasks. *Instructional Science*, 42(2), 271-290. doi: 10.1007/s11251-013-9280-7
- Beckmann, J.F., & Guthke, J. (1995). Complex problem solving, intelligence, and learning ability. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European Perspective* (pp. 177–200). Hillsdale, NJ: Erlbaum. doi: 10.4324/9781315806723
- Beckmann, J.F. (2018). Deferential trespassing: Looking through and at an intersectional lens. *New Directions for Child and Adolescent Development*, 161, 119-123. doi: 10.1002/cad.20243
- Beckmann, J.F., & Goode, N. (2017). Missing the wood for the wrong trees: On the difficulty of defining the complexity of complex problem solving scenarios. *Journal of Intelligence*, 5, 15. doi: 10.3390/jintelligence5020015
- Beckmann, J.F., Birney, D.P. & Goode, N. (2017). Beyond Psychometrics: The difference between difficult problem solving and complex problem solving. *Frontiers in Psychology*, 8, 1739. doi: 10.3389/fpsyg.2017.01739
- Birney, D.P., Beckmann, J.F., & Seah, Y.Z. (2016). More than the eye of the beholder: The interplay of person, task and situation factors in evaluative judgments of creativity. *Learning and Individual Differences*, 51, 400-408. doi: 10.1016/j.lindif.2015.07.007
- Borsboom, D., Mellenbergh, G.J., & Van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061-1071. doi: 10.1037/0033-295x.111.4.1061
- Dörner, D. & Funke, J. (2017). Complex Problem Solving: What It Is and What It Is Not. *Frontiers in Psychology*, 8, 1153. doi: 10.3389/fpsyg.2017.01153
- Gadanne, V. (1976). *Die Gültigkeit psychologischer Untersuchungen* [The validity of psychological inquiry]. Stuttgart: Kohlhammer.
- Lewin, K. (1927). Gesetz und Experiment in der Psychologie [Law and experiment in psychology]. *Symposium*, 1, 375-421.
- Lewin, K. (1992). Law and Experiment in Psychology. *Science in Context*, 5, 385-416. doi:10.1017/s0269889700001241
- Mencken, H.L. (1921). *Prejudices. Second Series*. London: Jonathan Cape.
- Wood, R.E., Cogin, J., & Beckmann, J.F. (2009) *Managerial Problem Solving: Frameworks, Tools & Techniques*. McGraw Hill Australia.

Corresponding author: Jens F. Beckmann, School of Education, Durham University, Durham DH1 1TA, United Kingdom. E-mail: j.beckmann@durham.ac.uk